

BAB III

METODOLOGI PENELITIAN

3.1 Metodologi Penelitian

1. Studi literatur

Pada tahap ini dilakukan pemahaman alur yang digunakan pada penelitian yang akan dilakukan. Literatur yang menjadi acuan adalah buku dan jurnal yang membahas mengenai pengolahan citra, klasterisasi data mining, dan algoritma K – Means.

2. Mendapatkan dataset

Pada tahap ini, data yang digunakan adalah data dari penelitian sebelumnya yang dilakukan oleh Agus Eko, tahun 2014 berupa data corel.

3. Implementasi algoritma

Pada tahap ini dilakukan proses pengolahan citra dengan menggunakan metode MTCD dan proses klasterisasi menggunakan metode K – Means. Metode MTCD digunakan untuk mengekstraksi ciri apa saja yang terdapat pada citra. Pada penelitian ini menggunakan 3 jenis ekstraksi ciri, yaitu ciri warna, tepi dan sudut. Metode K – Means digunakan untuk mengelompokan atau klasterisasi citra yang telah diekstraksi cirinya kedalam klasternya masing-masing.

4. Pengujian

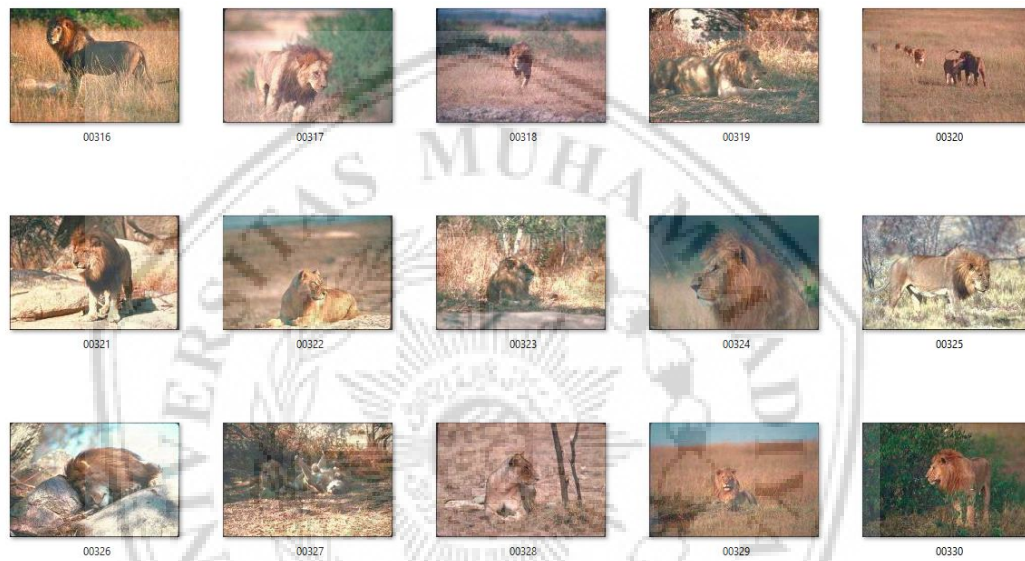
Pengujian yang dilakukan pada tahap ini dibagi menjadi 2, yaitu pengujian terhadap ketepatan atau presisi dari temu kembali citra dan pengujian terhadap nilai klaster yang paling optimal untuk digunakan. Pertama, untuk mendapatkan nilai presisi dari temu kembali citra menggunakan data uji dan dihitung menggunakan persamaan presisi. Kedua, untuk mendapatkan nilai klaster yang paling optimal digunakan pada penelitian ini adalah dengan menggunakan persamaan silhouette coefficient.

5. Penarikan kesimpulan

Pada tahap ini dilakukan penarikan kesimpulan pada penelitian yang telah dilakukan, apakah sesuai dengan yang diinginkan atau tidak

3.2 Dataset

Data yang digunakan adalah citra corel. Data citra diperoleh dari penelitian yang sudah dilakukan sebelumnya oleh Agus Eko, tahun 2014. Citra corel berjumlah 10.000 citra dengan jumlah 100 kelas dan masing-masing kelas berjumlah 100 citra. Pada gambar 3.1 merupakan contoh citra dengan kelas Singa dengan jumlah 15 citra. Citra dengan kelas singa merupakan citra yang di dalamnya terdapat objek singa.



Gambar 3.1 Contoh citra corel

Pada citra corel terdapat 100 kelas yang berbeda, gambar 3.1 merupakan contoh 1 kelas dari 100 kelas yang ada. Berikut merupakan daftar kelas pada dataset corel yang akan dijabarkan pada tabel 3.1:

Tabel 3.1 Daftar kelas dataset

No.	Kelas
1	Jendela
2	Beruang putih
3	Serigala
4	Singa

5	Gajah
6	Harimau
7	Gunung salju
8	Laki-laki berenang
9	Bangunan mesir
10	Sayur-sayuran
11	Model wanita
12	Anjing
13	Awan
14	Jamur
15	Kastil
16	Lukisan anak-anak
17	Suku orang dalam
18	Gelas
19	Laki – laki Silat
20	Galaksi
21	Album wanita
22	Buah – buahan
23	Tanaman dan pot
24	Plang toko
25	Kembang api
26	Pintu dan tangga
27	Rusa
28	Mobil balap
29	Jembatan
30	Air terjun
31	Padang rumput
32	Ruang tamu
33	Lukisan jepang
34	Plang jalan
35	Laut

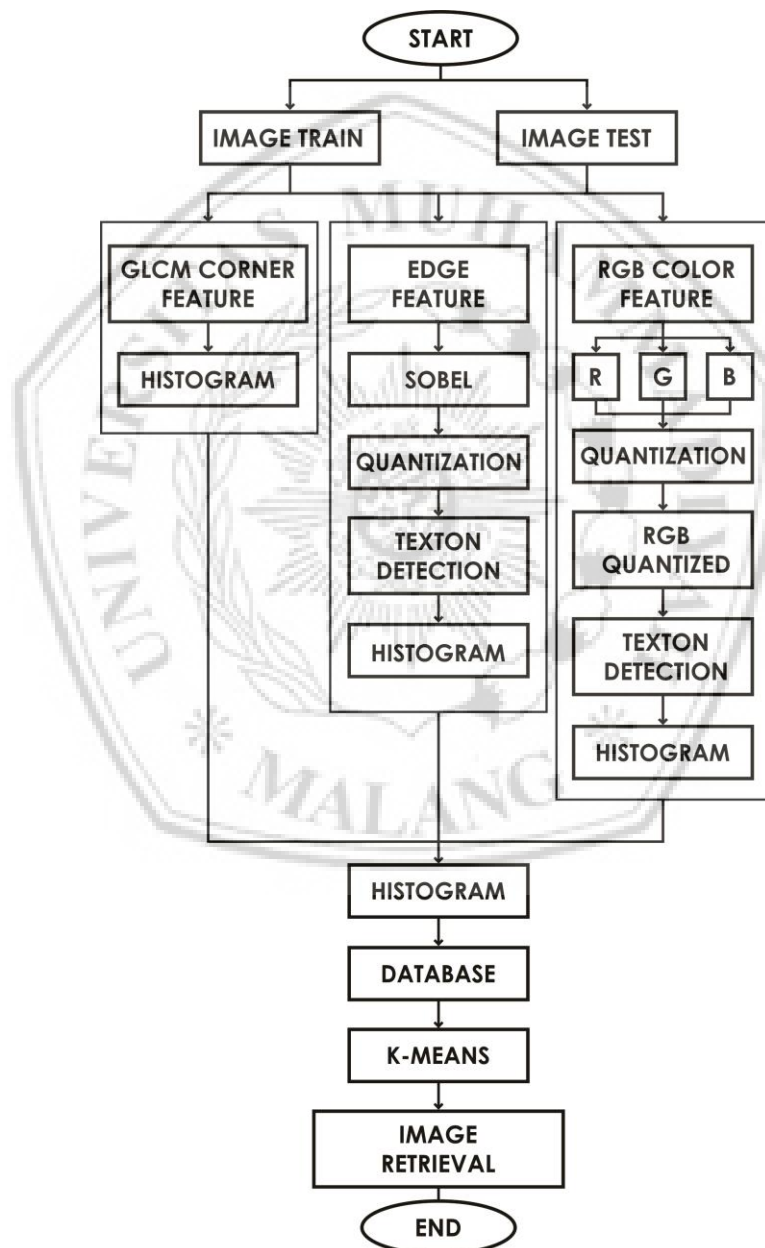
36	Kerikil
37	Matahari terbenam
38	Mobil klasik
39	Coretan dinding
40	Kartu remi
41	Kudani
42	Kupu – kupu
43	Teko
44	Boneka
45	Skrup
46	Rumah
47	Pistol
48	Bendera
49	Kereta klasik
50	Motif bunga
51	Petenis
52	Kartu tarot
53	Balon udara
54	Obat pil
55	Kapan pesiar
56	Makanan
57	Bus
58	Aurora
59	Telur jepang
60	Motor balap
61	Permen
62	Dinosaur
63	Tekstur sel
64	Sampul majalah
65	Kapal balap
66	Gurun pasir

67	Sel darah
68	Lukisan keramik
69	Pesawat
70	Meja dan kursi
71	Monyet
72	Domba
73	Baris – berbaris
74	Ombak
75	Kucing
76	Orang menggunakan pakaian renang
77	Lumut
78	Kristal
79	Hutan
80	Ikan
81	Motif batu
82	Motif bunga
83	Mercusuar
84	Elang
85	Batu
86	Cat warna
87	Berkuda
88	Benih
89	Manisan
90	Taman bunga
91	Kuda
92	Alat musik
92	Bunga
94	Daun
95	Angsa
96	Kalkun
97	Motif pakaian

98	Wanita berjemur
99	Kereta listrik
100	Lampu

3.3 Perencanaan Pengolahan Citra

Pada bagian ini akan menjelaskan mengenai langkah – langkah pengolahan citra yang dijabarkan pada gambar 3.2 dan uraian singkat dari tiap langkahnya.



Gambar 3.2 *Flowchart* perencanaan pengolahan citra

Proses awal dalam pengolahan citra yaitu melakukan pembagian data menjadi data latih dan data uji. Pembagian ini dimaksudkan untuk menggunakan data uji sebagai bahan untuk pengujian. Fungsi lain data uji adalah untuk dilakukan klasterisasi dimana telah ditambah dengan data latih. Setelah itu pada masing-masing data uji dan data latih dilakukan perhitungan menjadi tiga fitur yaitu warna, tepi, dan sudut. Berikut merupakan penjelasan masing-masing fitur:

3.3.1 RGB Color Feature

Warna pada masing-masing citra dibagi menjadi tiga penyusun warna, yaitu Merah(R), Hijau(G), dan Biru(B). Setelah warna dibagi menjadi tiga bagian, dilakukan kuantisasi pada masing-masing citra agar diperoleh rentang intensitas yang lebih kecil. Lalu hasil dari masing-masing kuantisasi digabungkan kembali menjadi satu kesatuan dan dilakukan deteksi texton. Kemudian hasil dari deteksi texton yang didapatkan dari deteksi 2x2 piksel horizontal atas, vertikal kiri, diagonal kiri, diagonal kanan, horizontal bawah, dan vertikal kanan.

3.3.2 Edge Feature

Melakukan perhitungan menggunakan Sobel untuk mendeteksi tepi dari citra. Setelah diperoleh hasil dari deteksi tepi, lalu dilakukan kuantisasi dan deteksi texton.

3.3.3 Gray Level Co-occurrence Matrix (GLCM) Corner Feature

Proses ini dilakukan untuk memperoleh informasi dengan mencari derajat keabuan setiap dua buah piksel yang dipisahkan dengan jarak. Derajat yang dicari yaitu 0° , 45° , 90° , dan 135° .

3.3.4 Histogram

Pada proses ini, data citra yang telah diekstraksi ciri kemudian direpresentasikan kedalam histogram, hal ini dilakukan untuk mempermudah ketika penulisan kedalam format .csv dalam bahasa pemrograman python.

3.3.5 Database

Data yang telah direpresentasikan kedalam histogram pada tahap sebelumnya kemudian ditulis dan diletakan dalam sebuah file dengan format .csv.

3.3.6 K – Means

Setelah melalui proses sebelumnya maka akan didapatkan hasil pada masing-masing fitur. Pada tahap ini akan diperoleh dua basis citra dari data uji dan data latih. Hasil dari pencocokan ukuran jarak kemudian dilakukan pembagian kategori dengan menggunakan K-Means sesuai dengan jumlah katogori yang diinginkan. dan akan diperoleh temu kembali citra yang telah dikategorikan.

3.3.7 Image Retrieval

Proses ini merupakan proses akhir dimana dilakukan pencarian citra terhadap dataset yang ada. Pertama, citra yang ingin dicari dimasukan kedalam program. Citra yang ingin dicari disebut dengan citra query. Setelah citra query dimasukan, maka program akan memproses dengan cara mengekstraksi ciri dari citra query dan dihitung kedekatan jaraknya terhadap citra yang telah di kelompokkan dalam kategorinya masing – masing pada proses K – Means. Setelah diketahui citra query tergolong dalam kategori mana, lalu dilakukan pengurutan sesuai dengan jarak yang paling dekat.

3.4 Lingkungan Kerja

Lingkungan kerja meliputi perangkat keras dan lunak.

3.4.1 Perangkat Keras

Menggunakan laptop merek ACER tipe E5 475G dengan spesifikasi perangkat keras yang akan dituliskan pada tabel 3.2 sebagai berikut:

Tabel 3.2 Spesifikasi perangkat keras

No	Spesifikasi
1.	<i>Processor</i> Intel Core i3-6006U CPU @ 2.00GHz
2.	RAM 4 GB DDR4L
3.	<i>Harddisk</i> 500 GB

3.4.2 Perangkat Lunak

Menggunakan perangkat lunak yang akan dituliskan pada tabel 3.3 sebagai berikut:

Tabel 3.3 Spesifikasi perangkat lunak

No	Perangkat Lunak
1.	Sistem Operasi Microsoft Windows 10 Ultimate 64-bit
2.	Jupyter Notebook
3.	Library OpenCV
4.	Ms. Word 2010

3.5 Evaluasi

Pada penelitian ini menggunakan metode *Silhouette Koefisien* untuk mendapatkan nilai evaluasi antara atribut satu dengan atribut yang lain. Berikut merupakan penjelasan mengenai evaluasi yang digunakan:

3.5.1 Silhouette Koefisien

Nilai *silhouette koefisien* memiliki macam antara -1 sampai 1. Dimana -1 artinya nilai klaster mengalami *overlapping* dan 1 sangat baik. Hasil klaster yang baik bernilai positif. Nilai klaster yang *overlapping* lebih tepat dimasukkan ke dalam klaster yang lain.

3.6 Pengujian

Dataset citra berjumlah 10.000 terdiri dari 100 kelas dan masing – masing kelas berjumlah 100 citra. Pada tahap pengujian ini menggunakan data uji berjumlah 5000 data. Data uji diambil dari 50 citra dari setiap kelas secara random. Tiap – tiap data citra di uji dan dihitung nilai rata – ratanya menggunakan persamaan *precision*.

3.7 Lini Waktu (Timeline)

Berikut merupakan rancangan waktu pengerjaan pada penelitian yang akan dilakukan, dijabarkan pada tabel 3.4 berikut :

Tabel 3.4 Lini Waktu (Timeline) Penelitian

Aktifitas	Waktu															
	Bulan 1				Bulan 2				Bulan 3				Bulan 4			
Minggu	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Menyusun Proposal																
Studi Pustaka																
Analisis																
Perancangan																
Implementasi																
Pengujian																
Perbaikan																
Penyusunan Laporan																